DOCUMENT RESUME

ED 305 100 JC 890 130

AUTHOR Licata, Paul; Snider, William J.
TITLE Minimum Competency Factors: In and Out of the Classroom.
INSTITUTION Valencia Community Coll., Orlando, Fla.
PUB DATE 87
NOTE 22p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Community Colleges; *Grades (Scholastic); Higher Education; *Minimum Competency Testing; *Predictor Variables; Required Courses; Scores; *Test Coaching; Two Year Colleges

ABSTRACT

In 1987, a study was conducted to evaluate the ability of several variables to predict student performance on Florida's minimum competency test for rising juniors, the College Level Academic Skills Test (CLAST). The variables included grades received in required courses; semesters elapsed prior to taking required courses after entering college; semesters elapsed between successful completion of coursework and taking the CLAST; completion of coursework in composition, reading, and math beyond required levels; and participation in a school-sponsored review session for one or more subtests of the CLAST. A random sample of 288 students selected from the 400 students from a moderately sized community college who took the 1984 CLAST was analyzed in terms students' status as first-time or repeat takers and in terms of students' participation in the review courses. Study findings included the following: (1) students with a grade of C or better in a class usually did well on that respective subtest of the CLAST; (2) among students who did not take a reading class, those who took the CLAST review course performed worse on the reading subtest than those who did not take the review course; (3) for students retaking the CLAST in 1984 without having taken a review course, grades in Composition I and II were the strongest indicators of success on the grammar subtest; and (4) the overall passing rate for the essay subtest was 85.6%, with 92.4% of those taking the test for the first time passing, compared to 27.2% of those retaking the test. (AAC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED305100

Minimum Competency Factors:

In and Out of the Classroom

Paul Licata

William J. Snider

ABSTRACT The purpose of this study was to evaluate the effects of using grade-in-class, lag time, and reviewing as a means of predicting MCT performance. Using a random sampling of previous student MCT scores and coursework performances, the researchers developed a robust statistical analysis of these variables, which hierarchically assigned them predictive values in a variety of testing situations. Results suggest that an analysis of academic class performance, timing, and reviewing may help to better counsel students in their separate approaches to MCT preparation.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P. Licata

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

IC 890 130

Copyright © 1987 by Paul Licata and William J. Snider.
All rights reserved.

BEST COPY AVAILABLE

Minimum Competency Factors: In and out of Classroom

The drive for educational accountability has, historically, been tied to the emergence of new testing strategies and philosophies. As Haney (1984) observed, Minimum Competency Testing is presently in vogue with both the general public and educational institutions. A sizable body of research has emerged to test the validity and reliability of Minimum Competency Testing instruments, its impact on educational policy making, and the demographic predictors of student performance. It is important to realize that once a state or school system institutes Minimum Competency Testing procedures, schools become bound by an implicit contract to graduate students with the skills to pass the Minimum Competency Test (MCT) (Cohen 1980), and it is on this criteria that students and schools are currently being evaluated. Studies of the Oklahoma Beginning Teacher's Exam and the MCT upon which this study will focus, The College Level Academic Skills Test (CLAST), have so far confirmed only that students with higher aptitude test scores (SAT, CAT, ACT) do better on MCTs than do students with lower scores on respective aptitude tests (Belcher 1985).

When evaluating predictors of success on minimum competency tests such as Florida's College Level Academic Skills Test (CLAST), the researchers found that conspicuously absent from the current literature on MCTs is an examination of the student's educational process at the institution purportedly preparing him or her for an MCT such as CLAST. Walstead (1984) points out that "...approaching the research problem from the perspective of the impact of specific variables on MCT performance, rather than of MCT performance on specific variables, may be more useful to education policy makers as they grapple with the MCT issue." Rather than looking at the impact of CLAST on institutional and educational policy making, this study looks to the student's course-taking history and its impact on the student's CLAST performance. Cross (1975) highlights the need to emphasize educationally relevant, rather than demographically convenient, variables. In addition to the traditional variables of aptitude test scores, race, and age, the present study included elements of each subject's education history: the sequencing of courses, success in those courses, the time lag between coursework and the CLAST examination, and previous CLAST experience. As McPhee and Kerr (1985) note that "...empirical studies of factors which may influence performance on proficiency examinations comprises a small part of the available literature," so this study also tests the impact of a student's attending a subtest-specific review session on his CLAST performance.

As this information was gathered through a review of student records, the populations were relatively unbiased in terms of demand characteristics and free of Hawthorne-effect concerns that might contaminate the data collected. Accordingly, a quasi-experimental design (Campbell and Stanley 1963) utilizing

---

The authors would like to note their appreciation to Dr. Elaine Greenwood, Dr. Emile Wasniewski, and the Department for Institutional Research for their generous assistance in this project.

naturally occurring between-group controls (those students who had chosen to attend or not attend a review session prior to taking the CLAST) was used to explore the following questions:

Q1. What general impact does reviewing have on MCT performance?

Q2. What are the specific circumstances which impact review efficacy?

Method

Subjects: The population (n=288) for this study was a randomized sampling of all (n=approx. 400) students who took the 1984 CLAST (Florida's minimum competency test for rising juniors) at a moderately sized Florida community college. The population for this study was separated into four categories: 1) First-taker/NR, those taking the exam for the first time who did not attend a review session prior to taking the exam; 2) First-takers/R, those who attended one or more of the free review sessions offered before testing; 3) Re-takers/NR, those re-taking the test who did not attend a review session before re-testing; and 4) Re-takers/R those re-taking the test who did attend one or more of the available review sessions before re-testing. All Re-take subtest data were included in the analyses (n=68).

Test Requirements: To pass the CLAST, a student needed to achieve a passing grade on each of the test's four subtests: Math (265), Reading (260), Grammmar (260), and Essay (4, holistically scored). Students in the population who either re-took only a specific subtest and/or only attended reviews pertinent to a specific subtest were accordingly re-coded into the appropriate category when that subtest was analyzed.

Design

A 2 x 2 Latin square (Re-takers/First-taker x Review/no Review) was generated for each of the four subtests.

Additional analyses of between-groups variation was done through a series of one-way analyses of variance and orthogonal t-tests. A stepwise multiple discriminant analysis (SMDA) was applied to each subpopulation to isolate the most predictive variable and/or the hierarchy of variables operating in each condition.

The selection of SMDA as the statistical model was predicated on methodological concerns about randomness violations in quasi-experimental designs. Powers' (1985) analysis of coached and not coached test-takers, a parallel to the review and no-review conditions of this study, strongly suggests that, as voluntary participants in such sections, the groups (coached vs not coached) are not equivalent populations and that regressive statistical procedures may be inappropriate in these non-

equivalent conditions. Fortune and Hutson (1984) note that when group differences on pre-test means exist, attempts to adjust data for fair comparisons may result in parcelling out part of the program effects. Further support for the use of stepwise procedures is the ability to avoid the distortion inherent in designs using a single dependent variable.

Powers (1985), notes that,

> ... unmeasured personal characteristics of coached and uncoached examinees have rendered the results of such studies (quasi-experimental) equivocal ... those who seek coaching may be highly motivated to perform well on the test (thus) coaching may appear to be effective even when it is not. Other unconsidered factors such as test anxiety, may relate positively to seeking coaching, but negatively to test performance, and may result in underestimates of the effect of coaching or even make it appear harmful (121).

Fortune and Hutson (1984) are concerned about quasi-experimental designs which, in an attempt to identify correlates of change, impose on the data a statistical model that alters the covariant structure. SMDA is used in the present study to isolate specific gain-factors without covariance manipulations, and to further quantify the effects of between-group differences. More important is the present study's application of SMDA to the re-test scores which controls for any personal differences, since the smae individual took both tests. By further dividing the subjects into groups who either did review or did not review prior to their re-testing, this study attempted to generate a more accurate measure of review efficacy.

Fortune and Hutson (1984) go on to warn that randomly equivalent control groups are usually not available in quasi-experimental research. The lack of ideal control generates two conditions which serve to confound the measurement of change. These adverse conditions are: 1) differences between groups on one or more of the unmeasured correlates of the criteria measured or 2) differences between groups on the pretest measures of criterion. To this concern, Fortune and Hutson note, "A simple t-test for correlated distributions is perhaps the most appropriate analytical model for studies which use their subjects as their own control. Those studies, however, are subject to erroneous assumptions about the characteristics of the subjects and about the influence of external forces." As the present study does use its own subjects in the control groups, this study uses the t-statistic not only to test for within-group changes but also to test for between group consistency. The creation of sub-population within this study based on the Test Experience x Review interactions is designed to minimize uncontrolled between-group factors. There is still the possibility that statistical artifacts could exist in other unmeasured variables. To this point, Fortune and Hutson (1981)

note that "as it appears now, no model or classification system is superior across all cases." They go on to suggest that the best model may be the one "least riddled with troublesome artifacts and most appropriate for a given situation." Since the present study seeks to understand the institutional factors which predict success on the CLAST (Florida's minimum competency test) for a very general population, individual personality differences are not a central concern here. The major concern is how the institution, through class offerings and counselling suggestions, can best provide its students with an educational experience base sufficient to facilitate MCT (CLAST) success.

TABLE 1 summarizes the 29 independant variables coded for each student in this study:

Grade-in-class	grade received in required courses: Composition I and Composition II (COMP I and COMP II), Reading 1105, and Math 1033
Delay	semesters elapsed prior to taking course-work after entering college
Lag	semesters elapsed between successful completion of coursework and taking CLAST
Other course-work	completion of coursework in the above mentioned disciplines beyond the required levels
Review	participation in a school sponsored review session for one or more subtests
Score [MCT (CLAST)]	score on CLAST subtests: Grammar, Essay, Reading, Math

Demographic Controls

Previous efforts in the area of MCT research illustrate the need for active control across important demographics. Accepting Gagne's (1985) identification of an interaction between familiarity and the retrievability of information, there are in the present study four distinct measures used to assess a student's prior familiarity with the material: his grade-in-class, other relevant coursework taken, whether or not he had

attended a review session prior to his taking the CLAST, and whether or not he had previously taken this exam. Waldstadt (1984) found that pre-testing emerged as an effective factor for predicting test scores, but not all forms of pre-testing were significant. The present study examines two types of pre-test sensitization that appear to strongly effect scores: reviewing and past test experience. Thus, for First-takers with no review, no pre-test sensitivity would exist; for First-takers who did review, there is the impact of the review alone; for Re-takers with no review experience, there is the experience gained only by having taken the test previously; and the Re-takers who did review would be the most sensitized, having both the review and the previous testing experience. Thus, a gradient of sensitivity is created within the latin square to enable accurate isolation of this factor for control in subsequent analyses.

To insure fair comparisons between groups, Shelly (1984) warns that the progress of a deficient skills group cannot be fairly measured when a skilled group is used as a control. By comparing only equivalent groups: Re-takers versus Re-takers, and First-takers versus First-takers, this study controlled for initial skill level differences. Again, t-tests performed on the independent variables in each comparison provide a quantitative measure of between-groups parallelism.

Weiner (1979, 1980, 1983), and Wolf and Savikas (1985) have provided insight into the role that "Time" plays in achievement on recall tasks. Gettinger (1985) operationalized Time as "time needed to learn a task," and found that different students need different amounts of time to achieve mastery. He also found that levels of initial learning and one-week retention drop significantly when students spent less time or were given fewer trials to learn the experimental task. This suggests that since the duration of a class is constant for all students, some will have had ample time achieve mastery of the material while others will not have had enough time to achieve equal mastery. The present study, in this regard, analyzes the ability of the review session to provide extra time for mastery-learning and improved test performance.

In Kapinus' (1985) study of factors facilitating recall performance, he found that "The only main effect found to be significant was the within-groups effect of test time . . . such that the mean score on immediate testing (6.70) was higher than the mean score for delay testing (5.92)." This documentation of the erosive effect of even a one-week lag between a student's finishing his coursework and his taking of CLAST would adversely affect the student's performance/score. Contrary to Belcher's (1985) suggestion that time plays only a secondary role in test performance, Kapinus seems to be arguing that any evaluation of minimum competency testing performance, such as on Florida's CLAST, must be sensitive to time as a primary variable. The present study examines the time variable in two forms: 1) delay: semesters elapsed between the student's enrollment and his taking a course; and 2) lag: semesters elapsed between coursework and

CLAST. In both forms Time is analyzed as a main effect and as an interactive variable with grad-in-class and reviewing as a predictor of CLAST success.

Results/Analysis

Grammar Subtest

Re-take Analysis:

The data from the non-reviewing Re-takers highlight factors which are associated with student success on the Grammar subtest of CLAST. The SMDA suggests that for the Re-taking/NR student, grades in Composition I and II are the strongest indicators of success, as these are the only variables which produce a significant change in Rao's V (Comp I grade V = 4.80, p = .0285; Comp II grade V = 7.82, p = .006). In fact, students in this group have a mean Comp I grade of 2.20 (four-point system with A=4) and Comp II grade of 2.29. The change in Rao's V effected by grades is consistent with Belcher's (1985) finding that "Time" between class completion and CLAST examination plays only a secondary role in predicting CLAST success. However, further inspection of the data suggests an interaction within this population. Both the SMDA and the Standardized Canonical Discriminant Function Coefficients identify the student's Comp I grade as the primary predictor, and the Comp II grade as the secondary predictor of the student's examination performance. But, the Standardized Discriminant Function establishes the lag time between the student's completion of Comp II and his re-testing as having approximately four times more predictive power than the student's grade in class (-.78 to -.19).

It may be that when these variables are analyzed within the stepwise procedure, the emphasis is on maximizing the explanatory power of the variables in combination. From this analysis, Comp II lag is relatively powerless from a "value-added" perspective. The final change in Rao's V associated with including each variable into the predictive group demonstrated that more variance is explained by the Comp II grade ($^V = 7.819, p = .005) than by Comp II lag ($^V = 2.194, p = .138). This hierarchy is also supported by the findings at step two in the stepwise analysis. After Comp I grade was selected as the most predictive, the predictive power of the remaining variables was determined. The V statistic for Comp II grade was 12.619 (F to enter = 4.276) as compared to Comp II lag's V of 10.067 (F to enter = 4.276). The inclusion of Comp II grade at step two added significant explanatory power to the hierarchy (F = 5.25, p = LT .036; V = 12.619, p = .0018). Thus the results of the SMDA for students in the Re-test/Review condition are more clear-cut. The lag time between Comp I and the test and the lag time between Comp II and the test listed these two variables, in this order, as most predictive (see Table 2). The order effect, as

determined from the SMDA is significant (final Wilk's Lambda =.0294, p = .0167), as is the Canonical Discriminant Function test, according to which this order explains 98.5 percent of the variance (Canonical Correlation = .985, X2= 15.868, p = .007). Only the delay in taking Comp I did not contribute significantly more power to the hierarchy when included (^V = .1526, p = .6960; Standardized Canonical Coefficient = .0329).

The data suggest that only two variables make a significant contribution to the explanatory power in this condition: the lag times between Comp I and Comp II, and the taking of the test. The stepwise addition of Comp II lag further reduced Wilk's Lambda from .10000 to .03134, (p = .00009) and increased Rao's V from 153.3 to 216.3 (out of a cumulative total of 231.0), thereby demonstrating that these lag factors are the primary vehicles for indicating performance in this condition. The addition of Comp I grade (^V = 9.55, p =.002) and Comp II grade(^V = 4.941, p = .026) significantly reduced the remaining unexplained variance. An inspection of the Lambda statistic after each step reveals that there was only 1.5 percent of the variance left unexplained before their inclusion into the model. Given the dramatically high level of significance after introducing the lag time of Comp II to the test, the inclusion of Comp I grade and Comp II grade actually eroded the predictive confidence from p = .00009 to p = .0025.

The standardized Canonical Discriminant Function Coefficients also suggest this delineation, as the coefficients for Comp I lag time (1.845) and the Comp II lag (-1.382) are very high when compared to Comp I grade (-0.381) and Comp II grade (0.252). Each time/lag variable has at least three times more predictive validity than either the Comp I or Comp II grade variable alone.

Table 2

Re-take Success Rate

	Review	No Review
Took Class	83.3%(5)	63.6%(7)
No Class	60.0%(6)	76.5%(13)

First-take Analysis

There was an extremely high passing rate for all First-

takers (97.75%) on this subtest. Non-reviewers passed at a rate of 98.4% (129/131) and reviewers passed at a rate of 95.7% (54/47). The t-tests revealed no significant differences between the groups on control variables, Comp I grade, Comp I delay, Comp I lag, Comp II grade, Comp II delay, Comp II lag, and Other coursework.

Although the reviewing group had lower scores on this section, the difference was not significant (299.17 to 310.67; t = 1.6, p = .11). Even within the context of high scores and high passing rates, there is some tacit support for taking a pretest review. In general, non-reviewers had higher grades and shorter lag times, a condition which this research suggest would predict success without reviewing. Reviewers, with their lower grades and longer lags, were in a relatively skill-deficient initial position and would be expected to score lower on the test. Reviewing seems to minimize the impact of the longer lag, as a 15.3 percent longer Comp II lag time translated into only a 3.7 percent lower score. Reviewing helps when there is a Comp I lag time of three or more semesters and/or a Comp II lag time of two or more semesters. Reviewing under short lag conditions actually leads to lower scores than those of their non-reviewing counterparts. A similar analysis conducted for Comp I and II grades found reviewing helpful only when coupled with a Comp I grade of C or below.

The very high passing rate (174 of 178) renders any dichotomous (pass/fail) analysis meaningless and makes the application of stepwise procedures inappropriate.

Essay Subtest Analysis

Re-take Analysis

The overall passing rate for the Essay subtest of CLAST was 85.6 percent, but there was a great disparity between those taking the test for the first time (92.4% passed) and those re-taking the test (27.2% passed). Although re-takers in general were not very successful, those who had reviewed fared better (37.4% passed) than their non-reviewing counterparts (only 21.4% passed).

Since a vast majority of first-takers pass this subtest, the re-take population is very small (non-reviewing 14, reviewing 8), and one should be very conservative about generalizing from such a small sample. The demographics of these two groups were remarkably similar. There were no differences between these two groups of Re-takers in terms of Comp I grade (2.0) to 2.0; p = 1), Comp I lag time (2.61 to 1.88, p = .871), Comp II grade (2.22 to 2.5; p = .55), Comp II lag (3.2 to 4.75, p = .22), Other related coursework (1.93 to 1.88, p = .72) nor on their initial essay score (2.0 to 2.25, p = .17). Given these parallel groups, the input of reviewing can be tested.

A scatterplot analysis of Success X Improvement showed a very strong relationship (r = .926, p = .00001) that explained

85.5 percent of the variance. Each person that improved (n = 6) upon his original score passed the re-take, and of that group 3 had reviewed and 3 had not. Of the ten subjects that did not exhibit any improvement, eight had not reviewed. For this small sample, a Fisher's exact probability of p= .254 suggests that reviewing prior to a re-take provides a positive but statistically insignificant influence on the re-take score. This may imply that reviewing operates more from an avoidance of failure motive than as an assurance of success. More research needs to be conducted to fully understand the differences observed here to determine if they are an artifact of the small sample size, the result of another confounding variable such as test anxiety or a true measure of the review's positive effect on the essay CLAST score.

First-take Analysis

There were no significant differences between reviewing and non-reviewing First-takers in terms of their Comp I grades (2.57 to 2.46, p = .40), Comp I delay (1.66 to 1.59, p = .81) Comp II grades (2.68 to 2.83, p = .31) nor Other related coursework (1.89 to 1.94, p = .27). Reviewers had only slightly longer lag times in both Comp I lag (4.22 to 4.52, p = .097) and Comp II lag (2.95 to 3.46, p = .087) and had no significant difference in CLAST score (non-Reviewers scored higher than did those who attended a review (4.87 to 4.51, p = .128). The presence of marginal differences in control variables and a marginal difference in the dependent variable could somewhat cloud the interpretation. If one views the groups as parallel because there were no significant differences, the lack of significant differences in their CLAST scores must likewise be interpreted as having "no difference," proving the review ineffective. Yet, if one views the groups as "fairly different," then the review failed to compensate for the longer lags. The conclusion is the same for either a conservative or liberal reading of the data: reviewing did not significantly help students on this subtest, even when operating in the hypothetically optimal condition of long lag.

The First-taker/non-reviewing group initially tested the ability of the variable to predict success (passing) on this subtest of the CLAST. A student's passing of this subtest was moderately correlated to his Comp I grade (x^2=24.56, p = .0004). No other variable approached significance. The population was sufficiently large (n=137) to apply SMDA for further clarification. This stepwise procedure showed a substantial amount of unexplained variance (Lambda = .91, p = .04) and suggests that other extraneous variables are operating in this condition. Of the variance explained by the factors tested, the hierarchy of Comp II lag, Comp I grade, and Comp I delay was most predictive (V = 11.42, p = .01). Neither Comp I grade ($^{\wedge}$V = 1.22, p = .27) nor Comp I lag ($^{\wedge}$V = .01, p = .93) added any power to the hierarchy. The Standardized Canonical Coefficients isolate Comp II grade (SCC = 1.05), Comp I delay (SCC = 1.00) and

Comp I grade (SCC = .89) as individual explanatory elements. The inclusion of Comp I delay among the predictive variables was surprising in light of its weak Chi-square effect. Examination of the correlation matrix gives some insight into this, as Comp I delay is most highly correlated with Comp I grade (r=.64) and shares its predictive power. An analysis of raw scores provides more precision than the dichotomous pass/fail configuration and lends more credibility to the power of Comp I grade. The student's Comp I grade had the strongest correlation (r=.27) with his score. His Comp II lag was more strongly correlated to his Comp I grade (r=.31) than to his score (r=.21) and his Comp I delay was virtually unrelated to score (r=.02) A scatterplot analysis of these variables yields but one significant finding, that of Comp I grade X score (r=.40, p = .00001). Thus, for the First-takers who did not review, Comp I grade is the most important variable to consider when attempting to predict success on the essay subtest.

The First-taker/Review population was studied in a similar manner. A within-groups analysis based on a pass/fail criterion failed to yield any significant findings; although Comp I delay (x^2=10.49, p = .06) and Comp II grade (x^2= 5.55, p = .14) did approach significance. The moderate sample size of First-takers who did not review allows for application of a SMDA, though caution should be exercised when extrapolating from so small a group to other populations. Subjecting this subtest data to the stepwise procedure indicated that the factors studied had sufficient explanatory power (Lambda = .86, p = .26) and selected the combination of Comp 1 delay and Comp II grade (V = 6.70, p = .04) as most predictive. These same factors had the highest Standardized Canonical Coefficients (Comp I delay = .86, Comp II grade =.73) while no other factor was greater than 0.41. This combination of predictive factors suggests that while a short Comp I delay and a high Comp II grade may predict success, neither is effective in predicting the actual score on the exam (Comp I delay X score r=.05; Comp II grade X Score r=.17). The factors with the highest correlation to Score a Comp II lag (r=.38) and Comp I lag (r=.26). The emergence of these new factors is due to their with-group correlations. Comp I delay is most highly associated with Comp I lag (r= -.44) and Comp II lag (r=. -.37) suggesting that a short delay means that the composition courses were taken early in the student's education program and consequently translate into longer lag time between the completion of his class and his taking the CLAST. Since there was no relationship between Comp I delay and Comp I grade (r=.02) nor between Comp I delay and Comp II grade (r=.04), it is safe to suggest that the effects of Comp I delay are subsumed by Comp I lag and Comp II lag. The other pass/fail predictor, Comp II grade, is also correlated with Comp I lag (r=.57) and Comp II lag (r=.35). Scatterplot analysis of these factors found two significant relationships: 1) Comp I lag X score (r=.29, r2=.084, p = .023; B=.19) and 2) Comp II lag X score (r=.30, r2=.09, p = .02; B=.33). Thus, the time/lag factor, in fact,

becomes the best predictor for success on this subtest for the Reviewing population.

The direct impact of reviewing is noted when comparing the within-group findings for First-takers with no review with First-takers who did review. When non-reviewing First-takers approach the test with only personal knowledge, Comp I grade becomes the primary predictor of success. Reviewing had a differential effect on students, as those with longer lag time scored higher than those with shorter lag times, even though their grades reflected no significant difference in their base of knowledge prior to the exam. From an applied perspective, students with higher grades in Comp I (C or above) and with a short lag time between the completion of Comp I and the taking of the exam may not benefit from a review, while students with longer lag times (3 or more semesters), regardless of grade, should be counselled to attend a review to maximize their scoring potential.

Reading Subtest Analysis

Re-take Analysis

There were no significant differences between the reviewing and non-reviewing retakers' grades (2.27 to 2.5, $p = .30$), lag time (3.6 to 3.9 semesters, $p = .465$) nor in the taking of Other reading courses before the CLAST (17.9% to 25%, $p = .30$). This common base provides for a fair test for the impact of a review session on the students' CLAST performances. As shown in Table 3, the most successful cell is the Review/class combination (83.3% pass), followed by No-Review/no class (76.5%), No-Review/class (63.6%) and finally, Review/no class (60%). This hierarchy does not achieve statistical significance (Fisher's exact test, $p = .25$), but it does imply that the Reading Review operates in a very traditional style, as it can enhance recall but is not effective as a teaching tool. Although the Review/class combination was most successful, it is equally important to note that the Review/no class cell is the least successful. Students who did not take the class and reviewed were 16.5 percent less likely to pass than their non-reviewing counterparts. Consistent with the current literature on interference and learning theory (Powers 1985), taking a review without having taken the course may actually be counterproductive.

The power of review is also seen in the improvement scores. The no-review group had a mean gain of 35.96 points when they re-took the exam, an increase attributable to test familiarity (sensitivity, practice effect, etc.). Members of the reviewing group had the same practice effects as their non-reviewing

counterparts plus a review session, and their scores increased by a mean of 41.78 points. The review session itself, in this sense, is "worth" approximately 6 points. This difference does not achieve statistical significance ($t = .49$, $p = .632$) but does have practical value in that, 22 of the 39 students who failed on their first attempt, missed the cutoff score by 6 or fewer points. A review conceivably could have reduced the first-failure rate by 56 percent.

The extremely low number of subjects in each of these populations (No-Review n=28, Review n=16) makes subsequent factor analytic testing (SMDA) inappropriate.

TABLE 3

Re-take Success Rate

	Review	No Review
Took Class	83.3% (5)	63%(7)
No Class	60.0%(6)	76%(13)

First-take Analysis

For these first-takers, both the reviewing non-reviewing groups were similar in their grade (2.79 to 3.17, $p = .11$), lag time (4.14 to 4.0 semesters, $p = .86$) and their taking of Other reading classes (12.3% to 20%, $p = .28$). The high success rates of these populations (93.9% to 97.1%) suggest that reviewing per se has little predictive value for these first-takers. The ability of the class to prepare marginal students for success on the CLAST is show in Table 4. A grade of "C" or better in the class is a strong predictor of CLAST success (Fisher's exact $p = .02$). All 6 of the review/class students passed the CLAST; 27 of the 28 (96.5%) who reviewed without taking the class also passed the CLAST. Students with high diagnostic reading scores were not required to take a reading class, as an appropriate level of performance already existed. This skill dimension is the only one that emerges from the within-group analysis, since time-lag is not predictive from either a graduated ($x2 = 3.69$, $p = .45$) or dichotomized (Table 5, Fisher's exact $p = .37$) analysis.

TABLE 4

CLASS/CLAST SUCCESS (for non-reviewers

		CLAST Pass	Fail
Class Grade	D or lower	0	2
	C or higher	24	3

TABLE 5 LAG/CLAST SUCCESS (for non-reviewers)

		Pass	Fail	
Lag Time	Short lag	5	0	(3 or more semesters)
	Long lag	24	5	(less than 3 semesters)

Only the non-reviewing population was of sufficient size (n=30) to justify application of the SMDA. Both the stepwise procedure and the Standard Canonical Coefficients establish grade-in-class as the most viable predictor of the group (dV = 13.17, p = .0003; SCC = 3.58) followed by the time factors of Lag(dV = 7.76, p = .005; SCC = 1.60) and Delay (dV = 4.38, p = .027; SCC = 1.47). The presence or absence of an advanced reading course had no impact on the predictive hierarchy (dV = .24, p = .63; SCC = 0.10). The significant Wilk's lambda (L = .831, p = .0001) indicates that much of the sample's variance is left unexplained and that an analysis of additional factors would be necessary for a full understanding of this population.

In sum, there is no evidence that taking a review either helps or hinders a student's performance on the CLAST reading subtest (t = .83, p = .41). Unlike the skill-deficient subjects in the re-taking condition, the decision to review in this condition be the student's own choice. While reviewing does not as significant content acquisition function, it may serve as a confidence builder

Math

Re-take Analysis

With 99.2 percent of the first-takers passing and 76.9 percent of the re-takers passing, the overall passing rate for this subtest was 97.5 percent. This leaves very few unsuccessful

cases available for analysis.

There were no differences between the reviewing (n=9) and non-reviewing (n=4) re-takers in their grades (p=.39), Lag time (p=.129) nor their taking of Other math courses (p=.35). Upon re-testing, the non-reviewers gained 39.3 points and reviewers gained 13.5 points, a significant difference (p=.046), even with the extremely small sample sizes. These findings may be an artifact of the small sample size because these differences did not translate into differential success rates (non-reviewers =77.7%, reviewers =75%) or into significantly different re-test scores (p=.39) even though their original scores were essentially equal (non-reviewers = 249.6, reviewers = 250.5; t = .18, p = .89). While further analysis can be little more than anecdotal, a pattern emerges suggesting that reviews work best in long-lag situations and that no review is needed for short-lag conditions. A test of the lag X review condition, however failed to reach significance (Fisher's exact p = .40).

First-take Analysis

There were no differences between the reviewing and non-reviewing First-takers in Grade (2.5 to 2.42, p = .73) nor in Lag time (3.26 to 3.0, p = .24). There was a slight difference (p = .08) between the groups in subsequent coursework, as 48 percent of the non-reviewers and only 33.3 percent of the reviews had taken an additional math class prior to the CLAST. This difference does not predict success on CLAST, for even though all 87 (100%) students with additional math courses passed this subtest, 106 of the 108 (98.1%) passed without having taken another math course. Likewise, reviewing had no impact on either success rates (98.7% to 100%) or raw scores (295.7 to 299.1, p = .31) on this CLAST subtest.

Since no between-group differences were noted, the analysis focused on any within-group configurations which may lend insight into the observed success profile. The sample size for the first-taker/no review condition (n=150) is large enough to warrant application of an SMDA, but the extremely distorted pass/fail ratio (148/2) precludes its use.

Inspection of the within-groups correlation matrix (Table 6) reveals that subsequent coursework has the highest correlation to CLAST score (r = .26) of any factor.

TABLE 6 Correlation Matrix
First-takers/No Review

	MATG	MATDLY	MATLAG	OTHRMATH	Math Score
MATG	1.000				
MAYDLY	.84098	1.000			
MATLAG	.85643	.70542	1.000		
OTHRMATH	.01580	-.06868	.10687	1.000	
Math Score	.15189	.07220	.13193	.26326	1.000

Correlation Matrix
First-takers/Review

TABLE 7

	MATG	MATDLY	MATLAG	OTHRMATH	Math Score
MATG	1.000				
MATDLY	.87471	1.000			
MATLAG	.83643	.63762	1.000		
OTHRMATH	-.17990	-.18926	-.06196	1.000	
Math Score	.04084	-.00869	-.01711	.17295	1.000

Since all reviewing first-takers passed this subtest, it is impossible to apply stepwise procedures. The within-group correlations for reviewers (Table 7) follow the same pattern as those of the non-reviewers: other coursework correlates most highly with CLAST score (r = .17) while Grade-in-class, Delay, and Lag factors are not predictive.

The scatterplot analysis for reviewers' Grade X Score shows a moderate correlation (r = .52) of good fit (p=.001) that exlains 27.3 percent (r2 = .273) of the variance in the population. The reviewer's grade explained almost three times as much variance os did the non-reviewer's grade-in-class. It seems that review helps students recall more of their class material, as the median score for non-reviewers was 291 and the median score for reviewers was 296.

Conclusions

The results of this study indicate that Grade, Lag-time, and Review all play a role in CLAST success. Students with a grade of C or better in a class usually did well on that respective subtest of the exam. However, if the student had a low grade, Lag-time and Review interacted to impact his or her score on that subtest.

The stepwise analysis exposed some grade-in-class/lag-time/review relationships which parallel the Classical Conditioning theorists' acquisition/extinction/spontaneous recovery model. If the relationship is found to be consistent, then testing through the acquisition/extinction(lag-time)/recovery(review or no review) model would seem worthwhile. For example, this paradigm explains the failure of review when uses as a the primary teaching tool. The reading subtest analysis showed that those students who did not take the course (thus, had no course-related acquisition) and then took a review (cue for recovery) did not fare as well on the test as did those students who took no review. On the other hand, students who did not review had no new unprocessed and half-learned material to confuse their intuition, and so scored higher on that section of the test. Perhaps better prepartation for MCTs will result when we understand how course-taking history conforms to the acquisition/extinction/recovery construct.

There can be no substitute for a well prepared student when achievement scores are being considered. To maximize the potential of the review, there must by an identification of any approach-avoidance characteristics of the review session which may inadvertently limit attendance. It was evident, for example, in the Reading First-taker/no review group that something other than the course-taking variables was responsible for student success. This may be attributable to such factors as test anxiety and low goal expectancy (Wolf and Savickas, 1985). Although this study does not address these issues directly, it does seem safe to suggest that to maximize student performance, an analysis of student anxiety levels and the consistency between review information presented and actual test demands is paramount to an understanding of how reviews affect scores. An examination of the lag-time from review session to test date may help us to help our students optimize their application of the recalled information, as even a lag of one week can affect quality of recall (Haney, 1984). Controlling for the course-taking history, the student's grade-in-class and the counselling of students into or out of a review session should imporve student performance on MCT where the minimal skills may be acquired at different periods in the student's educational career.

A review of the current data, research and theory suggests a need for continued study in minimal competency testing. As school systems move toward MCTs to evaluate achievement, so the schools will be expected to provide instruction which produces acceptably high measures of student learning. Cohen (1980) attributes a "production function" to the service provider, in this case, the educational institution. In the attempt to answer the research questions that guided the present study, some important insights were gained into the complex nature of minimum competency testing which should provided fertile ground for future hypothesis generation and testing. Replications are needed which include larger and/or weighted sampling techniques if more is to be learned about the pass/fail dimension of MCTs. A practical summation of the results is provided in appendix A to serve as a guide to counsellors, researchers and educators concerned with maximizing student performance on minimum competency tests.

To appreciate these responsibilities, we need only remember what success or failure on these tests means to students. In an interview with the Miami Herald, a twenty-two year old Miami Dade student may have echoed the perception about the CLAST, and perhaps MCTs in general, when she noted, "You could say your life depended on it" (Lopez 87).

Appendix A

Counselling Suggestions

RETAKERS

	Grade in Class	Long Lag (3+ sems)	Short Lag (less than 3 sems
WRITING	C or above	No Review	No Review (optional)
	below a C	Review	Review
ESSAY	C or above	Review	Review
	below a C	Review	Review
READING	C or above	Review	Review
	below a C (If no class)	Review (no review)	Review (no review)
MATH	C or above	Review	Optional
	below a C	Review	Optional

FIRST TAKERS

	Grade in Class	Long Lag (3+ sems)	Short Lag (less than 3 sems)
WRITING	C or above	Review	No Review
	below a C	Review	Review
ESSAY	C or above	Review	No Review
	below a C	Review	Review
READING	C or above	No Review	No Review
	below a C	Optional	Optional
MATH	C or above	Review (or another class)	Review (or another class
	below a C	Review (or another class)	Review (or another class

Works Consulted

Chase, J. L., and Chase, Richard B. (1976). Short notes: A statistical power analysis of applied psychological reasearch. Journal of Applied Psychology, 61 (No. 2), 234-236.

Babbie, Earl R. (1973). Survey research methods. California: Wadsworth Publishers Co., Inc.

Belcher, Marcia J. (1985). "Predicting clast perfomance: The roles of entering and basik skills and curriculum." 18th Annual Florida Statewide Conference on Institutional Research.

Campbell, Donald T. & Stanley, Julian C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

Cross, Patricia K. (1983). The state of the art in needs assessments. Community/Junior College Quarterly, 7. 195-206.

Davey, Beth and Kapinus, Barbara, A. (1985). Prior knowledge and recall of unfamiliar information: Reader and text factors. Journal of Educational Research, 7 (No. 3) 147-151.

Draper, Norman R. & Willard, Lawrence E. (1970). Probability: An introductory course. Chicago: Markham Publishing Company.

Fagley, N. S. (1985). Applied statistical power analysis and interpretation of nonsignificant results by research consumers. Journal of Counseling Psychology, 32 (No. 3) 391-396.

Fortune, Jim C. & Hutson, Barbara A. (1984). Selecting models for measuring change when true experimental conditions do not exist. Journal of Educational Research, 77 (No. 4) 197-205.

Fryer, Holly C. (1966). Concepts and methods of experimental statistics. Boston: Allyn & Bacon, Inc.

Gagne, Ellen D., Bell, Michael S., Yarbrough, Donald B., & Weideman, C. (1985). Does familiarity have an effect on recall independent of its effect of original learning. Journal of Educational Research, 79 (No. 1) 41-45.

Haney W. (Winter, 1984). Testing reasoning and reasoning about testing. Review of Educational Research, 597-654.

Harman, Harry H. (1967). Modern factor analysis. Chicago: The University of Chicago Press.

Gettinger, M. (1985). Time allocated and time spent relative to time needed for learing as determinants of achievement. Journal of Educational Psychology, 77 (No. 1) 3-11.

Lopez, Vivian. (1987). Quoted in "CLAST is: a)good; b)bad; c)required." Miami Herald, March 10, 1987, 1B.

Lyon, Mark A. and Gettinger, M. (1985). Differences in student performance on knowledge, comprehension, and application tasks: Implications for school learning. Journal of Educational Psychology, 77 (No. 1) 12-19.

McPhee, Sidney A. & Kerr, Michael E. (1985). Scholastic aptitude and achievement as predictors of performance on competency tests. Journal of Educational Research, 78 (No. 3) 186-190.

Mischel, Walter (1968). Personality and assesment. New York: John Wiley and Sons.

Owen, E. and Sweller, J. (1985). What do students learn while solving mathematics problems? Journal of Educational Psychology, 77 (No. 3) 272-284.

Palumbo, Denis J. (1969). Statitics in political and behavioral science. New York: Merridith Corp.

Powers, Donald E. (1985). Effects of coaching on GRE aptitude test scores. Journal of Educational Measurement, 22 (No. 2) 121-136.

Reynolds, G. S. (1968). A primer of operant conditioning. Illinois: Scot, Foresman.

Shelly, Joseph F. (1984). Evaluation of the centralized, structured, after-school tutorial. Journal of Educational Research, 77 (No. 4) 213-217.

Sindelar, Paul T. Gartland, D. & Wilson, R. J. (1984). The effects of lesson format on the acquisition of mathematical concepts by fourth graders. Journal of Educational Research, 78 (No. 1) 40-45.

Weiner, B. (1979). Title of art. Journal of Educational Psychology 75, 530.

---------- (1980). Human motivation. New York: Holt, Rhinhart, Winston.

ERIC Clearinghouse for Junior Colleges MAY 05 1989